



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2018

---

## **Serial recall of colors: Two models of memory for serial order applied to continuous visual stimuli**

Peteranderl, Sonja ; Oberauer, Klaus

**Abstract:** This study investigated the effects of serial position and temporal distinctiveness on serial recall of simple visual stimuli. Participants observed lists of five colors presented at varying, unpredictably ordered interitem intervals, and their task was to reproduce the colors in their order of presentation by selecting colors on a continuous-response scale. To control for the possibility of verbal labeling, articulatory suppression was required in one of two experimental sessions. The predictions were derived through simulation from two computational models of serial recall: SIMPLE represents the class of temporal-distinctiveness models, whereas SOB-CS represents event-based models. According to temporal-distinctiveness models, items that are temporally isolated within a list are recalled more accurately than items that are temporally crowded. In contrast, event-based models assume that the time intervals between items do not affect recall performance per se, although free time following an item can improve memory for that item because of extended time for the encoding. The experimental and the simulated data were fit to an interference measurement model to measure the tendency to confuse items with other items nearby on the list-the locality constraint-in people as well as in the models. The continuous-reproduction performance showed a pronounced primacy effect with no recency, as well as some evidence for transpositions obeying the locality constraint. Though not entirely conclusive, this evidence favors event-based models over a role for temporal distinctiveness. There was also a strong detrimental effect of articulatory suppression, suggesting that verbal codes can be used to support serial-order memory of simple visual stimuli.

DOI: <https://doi.org/10.3758/s13421-017-0741-0>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-139227>

Journal Article

Accepted Version

Originally published at:

Peteranderl, Sonja; Oberauer, Klaus (2018). Serial recall of colors: Two models of memory for serial order applied to continuous visual stimuli. *Memory Cognition*, 46(1):1-16.

DOI: <https://doi.org/10.3758/s13421-017-0741-0>

Serial Recall of Colors.

Two Models of Memory for Serial Order Applied to Continuous Visual Stimuli

Sonja Peteranderl

and

Klaus Oberauer

University of Zurich

Peteranderl, S., & Oberauer, K. (2018). Serial recall of colors: Two models of memory for serial order applied to continuous visual stimuli. *Memory & Cognition*, 46, 1-16.  
doi:10.3758/s13421-017-0741-0

The research reported in this article was supported by a grant from the Swiss National Science Foundation (grant number 100014\_126766) to the second author. Correspondence should be addressed to: Sonja Peteranderl, who is now at the ETH Zurich, Department for Education Sciences, Clausiusstr. 59, 8092 Zürich; email: sonja.peteranderl@ifv.gess.ethz.ch

## Abstract

This study investigates the effects of serial position and temporal distinctiveness on serial recall of simple visual stimuli. Participants observed lists of five colors presented at varying, unpredictably ordered inter-item intervals, and reproduced the colors in their order of presentation by selection on a continuous response scale. To control for the possibility of verbal labelling, articulatory suppression was required in one of two experimental sessions. Predictions were derived through simulation from two computational models of serial recall: SIMPLE represents the class of temporal distinctiveness models, whereas SOB-CS represents event-based models. According to temporal distinctiveness models, items that are temporally isolated within a list are recalled more accurately than items that are temporally crowded. In contrast, event-based models assume that time intervals between items do not affect recall performance *per se*, although free time following an item can improve memory for that item because of extended encoding. The experimental and the simulated data were fit with an interference measurement model to measure the tendency to confuse items with other items nearby on the list – the locality constraint – in people as well as in the models. Continuous-reproduction performance showed a pronounced primacy effect with no recency, and some evidence for transpositions obeying the locality constraint. Though not entirely conclusive, the evidence was leaning in favor of event-based models and against a role for temporal distinctiveness. There was a strong detrimental effect of articulatory suppression, suggesting that verbal codes can be used to support serial-order memory of simple visual stimuli.

## Serial Recall of Colors.

### Two Models of Memory for Serial Order Applied to Continuous Visual Stimuli

Maintaining and retrieving information in its correct order is essential for cognitive tasks. Simple tasks, such as remembering a telephone number, or more complex tasks, such as solving mathematical problems or writing an article, require serial order information. Over the past five decades much research in short-term and working memory has been devoted to the mechanisms of maintaining the serial order of a list of items. Several computational models aim to explain processes and mechanisms underlying memory for serial order (for a review see Lewandowsky & Farrell, 2008).

Forward serial recall has been studied extensively with verbal and spatial materials (for a review see Hurlstone, Hitch, & Baddeley, 2014), and also to some extent with complex visual stimuli such as novel visual patterns (Avons, 1998; Avons & Mason, 1999) and faces (Smyth, Hay, Hitch, & Horton, 2005; Ward, Avons, & Melling, 2005). In contrast, forward serial recall has so far not been studied with simple visual stimuli such as colors or orientations. There has been a surge of interest in working memory for simple visual stimuli in the past two decades (for reviews see Luck & Vogel, 2013; Ma, Husain, & Bays, 2014). The paradigms used in this literature rely nearly exclusively on the simultaneous presentation of memory items in a spatially distributed array. A few more recent studies have investigated immediate memory for sets of colors or orientations presented sequentially, followed by a single test of a randomly selected item (Gorgoraptis, Catalao, Bays, & Husain, 2011; Kool, Conway, & Turk-Browne, 2014). In these studies, the item to be tested was identified by a non-temporal feature such as its spatial location, so that participants did not have to remember the item's temporal order. Therefore, we know very little about how people remember the order of simple visual features presented sequentially. Whereas memory for serial order for verbal and spatial stimuli relies to a large extent on the same principles (Hurlstone et al., 2014), it is not clear whether these principles also generalize to simple visual stimuli.

With the present study we investigate whether the reproduction of visual stimuli in the order of presentation yields patterns of effect analogous to those found with verbal and spatial materials. We asked participants to remember lists of five colors, and to reproduce each color by selecting it on a color wheel. This paradigm combines the advantages of the continuous-reproduction (a.k.a. delayed-estimation) task with those of the classic serial-recall task. The continuous-reproduction task (Wilken & Ma, 2004) provides a measure of reproduction error on a continuous scale that can be used to measure the precision of memory representations (Ma, Husain, & Bays, 2014). The serial-recall task provides information about the effects of serial position and transposition errors. The serial-position curves for forward serial recall of verbal and spatial items is characterized by an extended primacy and a smaller recency effect. Transposition errors (i.e., confusions between list items) obey the "locality constraint": List items tend to be confused more often with other list items close to them in the list than with more

distant list items (Hurlstone, et al., 2014). Here we investigate whether the same patterns are observed in serial recall of visual features.

In addition, we tested for temporal isolation effects, because they help to adjudicate between time-based and event-based models of memory for serial order. These classes of models differ in what role time plays in maintaining and retrieving events from memory (Brown & Chater, 2001; Lewandowsky, Wright, & Brown, 2007). Time-based models assume that events are associated to the current state of a temporal context that gradually changes over time. In contrast, event-based models assume that events are associated to a context that merely represents their ordinal position in time (e.g., an item's ordinal list position). The aim of the present study is to provide a first test of these competing assumptions for the serial recall of visual items.

Time-based models assume a close relation between time and memory. A particular class of time-based models, known as temporal distinctiveness models, assume that memory for items in a list is determined by their temporal distinctiveness from other items in memory (Brown, Neath, & Chater, 2007; Burgess & Hitch, 1999, Glenberg & Swanson, 1986). These models assume that temporally isolated items – preceded and followed by relatively long time intervals – are remembered better than temporally crowded items. This effect is called temporal isolation effect. Temporal distinctiveness models also predict an event to be less retrievable as time increases since its encoding (Brown, et al., 2007). In contrast, event-based models assume no direct relation between time and memory. According to these models, time has a more indirect role, enabling cognitive processes that improve memory (e.g. Farrell & Lewandowsky, 2002, 2008b; Nairne, 1990; Oberauer, Lewandowsky, Farrell, Jarrold, & Greaves, 2012). Event-based models can predict that longer time intervals following a list item improve that item's memorability by assuming additional encoding or consolidation processes during that time, but they predict no beneficial effect of longer times preceding an item. In contrast, time-based models predict that both longer times preceding an item and longer times following it should render that item more temporally distinctive, and thereby better accessible in memory.

A considerable number of studies has been conducted to test the role of temporal isolation on memory for lists of mostly verbal, and occasionally spatial, stimuli. Temporal isolation effects have been observed in a variety of experimental paradigms, including recognition (Morin, Brown, & Lewandowsky, 2010), free recall (Brown, Morin, and Lewandowsky, 2006), free reconstruction of list order (Lewandowsky, Nimmo, & Brown, 2008), and running memory span (Geiger and Lewandowsky, 2008). In contrast, temporal isolation effects have been consistently absent in immediate forward recall. In the forward serial-recall task, a consecutively presented list of items has to be remembered and subsequently recalled in the presented order. Studies with this paradigm have shown that temporal isolation effects are absent in a broad range of circumstances (Farrell, Wise, & Lelièvre, 2011;

Lewandowsky, Brown, Wright, & Nimmo, 2006; Nimmo & Lewandowsky, 2005, 2006; Parmentier, King, & Dennis, 2006; Lewandowsky, Nimmo, and Brown, 2008; Saito, 2000).

### Computational Models of Serial Recall

Computational modelling is an indispensable tool to formulate and test theories quantitatively. With computational modelling, theories gain precision and are easier to falsify (Lewandowsky & Farrell, 2010). Models can be used in different ways to serve different purposes. One currently very popular use of models is to fit them to data, thereby optimizing the model parameters for the data at hand. This approach serves to assess how well a model can accommodate the data. Another way of using models is to generate predictions for a new experiment without fitting the model to the new data. This approach serves to assess which pattern of data the model predicts for the new experiment. Because the model parameters are not optimized, the predictions cannot be expected to be quantitatively as accurate as those after fitting the model. The advantage of this approach is that we can let the model make predictions for many experiments with different experimental paradigms and designs with no, or very few, changes to its parameters (for a cogent formulation of the rationale of this approach see Wills & Pothos, 2012). Here we follow this second approach.

We use two models for deriving predictions for our experiment through simulations: The SOB-CS model represents the class of event-based models, and SIMPLE represents the class of time-based models. We used each model with its standard parameter values, with two exceptions: We added a parameter to both models for the precision of color representations so that we can apply the models to the error data from the continuous-reproduction task. Moreover, for each model we adjusted one parameter by hand to approximate the overall level of accuracy in the data; for SOB-CS we adjusted the color-precision parameter for that purpose because it was a new parameter for which there are no standard values; for SIMPLE we used the *c* parameter translating temporal distance into similarity, because this parameter has been used in previous applications of SIMPLE to calibrate overall performance (Brown et al., 2007). Our aim was not to maximize the quantitative fit between model predictions and data but to investigate how well each of the models can predict the qualitative pattern of the data. We next give a brief introduction into these two models.

#### Event-based model: SOB-CS

SOB-CS (Oberauer, Lewandowsky, Farrell, Jarrold, & Greaves, 2012) is the latest instantiation of the SOB (“serial order in a box”) model (Farrell & Lewandowsky, 2002; Farrell, 2006; Lewandowsky & Farrell, 2008b).

SOB-CS is a two-layer neural network, in which one layer represents serial positions and one layer represents items. Memory for order is maintained by the associations between item representations and position representations (a.k.a. position markers). SOB-CS states that the encoding strength is based on an item's novelty. The more novel an incoming item, the more strongly it is encoded. This mechanism is called novelty-gated encoding. SOB-CS further assumes that working memory capacity is limited only by interference. Interference arises from superimposition of associations, and from confusion of different items dependent on their similarity.

One difference between SOB-CS and previous versions of SOB that is relevant in the present context is that SOB-CS explicitly models the duration of encoding: encoding strength increases as a function of the time that is spent to encode an item. Thereby, the model predicts a beneficial effect of the time interval following presentation of an item. That said, with its standard parameter settings encoding strength reaches an asymptote after about 500 ms, so that the effect of post-item time should be limited.

SOB-CS has been successfully applied to serial recall with discrete items such as letters, words, or digits, but has not yet been applied to stimuli varying on a continuous dimension, such as colors or line orientations. We therefore developed a new representational scheme suited for stimuli varying continuously in a low-dimensional feature space, presented below.

### **Temporal distinctiveness model: SIMPLE**

The SIMPLE (Scale Invariant Memory, Perception and LEarning) model is a formal instantiation of the temporal distinctiveness hypothesis. SIMPLE assumes that events are associated to a temporal context, and that context is an important – and often the only – retrieval cue through which the representation of a past event is accessed. The probability of accessing the searched-for event – for instance, the first item in the last-presented memory list – depends on that event's temporal distinctiveness at the time of retrieval. The confusability between items is related to the ratio of the time elapsed between their time of encoding and the retrieval time. The lower the ratio, the lower the confusability and hence the more likely it is that the item is correctly recalled (Brown, Neath, & Chater, 2007). As a consequence, items from further in the past, and items which are temporally crowded, are more difficult to recall.

Whereas the original SIMPLE contained only a temporal context, Lewandowsky, Duncan, and Brown (2004) extended the model by a second context dimension representing ordinal list position. The relative weight of the temporal dimension is controlled by a parameter *wt*. With a small weight on the temporal dimension SIMPLE predicts only negligible effects of temporal distinctiveness, which defies our purpose of using SIMPLE as a temporal-distinctiveness model. Therefore, we used the original

version of SIMPLE in which list items are discriminated only on a temporal context dimension (effectively setting  $w_t = 1$ ).

### Method: Experiment

We investigate the reproduction of a list of colors in serial order. By analogy to common findings in serial recall of other materials, we predict a serial-position curve with extended primacy and a small recency effect. Transposition errors should obey the locality constraint: To the degree that participants confuse a color with another color in the list, these confusions should tend to come from colors nearby in the list order. To test for temporal-isolation effects, we used the paradigm introduced by Lewandowsky, Nimmo, and Brown (2008): Lists of five items were presented with varying inter-item intervals. Four different inter-item intervals were assigned in a counterbalanced manner to the four inter-item gaps, thereby varying the temporal isolation of each list item – except for the first and the last -- independently of its serial position in the list. The sequence of inter-item intervals varies in an unpredictable way from trial to trial, so that participants cannot exploit a predictable temporal pattern for strategies such as grouping. With this paradigm, temporal isolation effects have repeatedly shown to be absent in forward serial recall of verbal and spatial items (Lewandowsky, et al., 2006; 2007; Nimmo & Lewandowsky, 2005, 2006; Parmentier et al., 2006). Based on these findings, we predict to find no effect of temporal isolation for lists of colors either.

Lewandowsky et al. (2007) argued that benefits of time on recall performance might arise from processes such as rehearsal, refreshing or consolidation, but only during the post-item interval (i.e., the interval directly after the presented stimulus), not the pre-item interval. With visual stimuli, a longer post-item interval could provide an opportunity to generate a verbal label of the just-seen color, which could aid memory. To control for this particular possibility, we tested participants for two sessions, of which one session involved articulatory suppression to minimize the chance of forming and rehearsing verbal representations of colors in post-item intervals.

### Participants

Twenty-one undergraduate students of the University of Zurich (3 male, 18 female) between 18 and 35 years participated in exchange for course credit. They completed two one-hour sessions.

### Materials and Procedure

The experiment was programmed with MATLAB using the Psychophysics Toolbox extension (Brainard, 1997; Pelli, 1997; Kleiner et al, 2007). Both sessions were run on identical 19" Dell monitors using an Asus EAH6450 graphic card to guarantee consistent presentation of colors in each trial. The



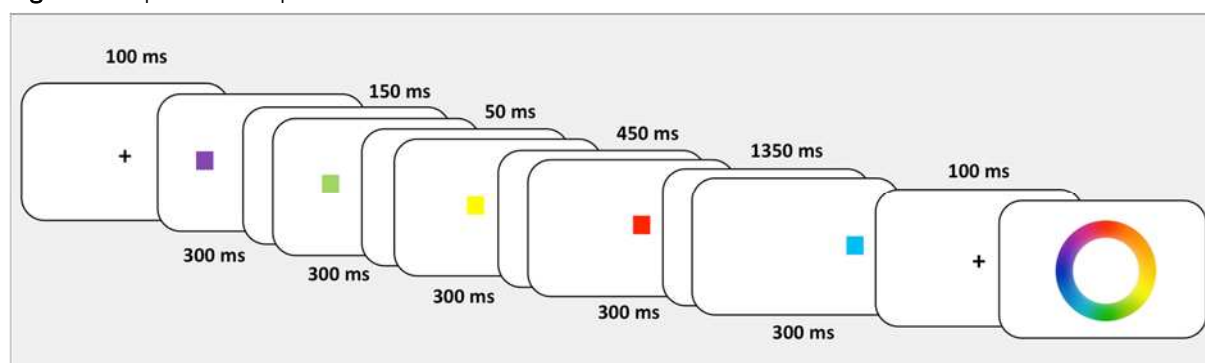
monitors had a refresh rate of 60 Hz and a resolution of 1280x1024. Color management settings were set to their default for the Windows 7 operating system.

The memory task was identical in both experimental sessions. Each session stands for one condition that differed in requiring articulatory suppression. In one condition, the participants did the experiment silently. In the other condition, articulatory suppression was required by speaking “bibibibi” aloud while doing the memory task. The order of conditions was counterbalanced across participants. The participants in the articulatory suppression session wore headsets and their voices were recorded every eighteenth trial to make sure that they continuously verbalize “bibibibi”. Participants started each session with four practice trials followed by 240 testing trials.

Figure 1 illustrates the sequence of events in one trial. Each trial began with a fixation cross in the middle of the screen for 100 ms. Afterwards, five colored squares (1 cm<sup>2</sup>) were presented consecutively, vertically central on the screen, and horizontally shifted from left to right to prevent masking effects. The first square started in the center of the left half of the screen, and each following square was shifted an equal distance to the right, such that the last square was presented in the center of the right half of the screen. Colors were randomly chosen from 360 colors of a color wheel. Each stimulus was presented for 300 ms and separated from the next stimulus by varied inter-item-intervals. These inter-item-intervals were either 50 ms, 150 ms, 450 ms, or 1350 ms long.

Each trial included all four intervals, and across all trials each inter-item-interval occurred an equal number of times in the same position. Hence the overall duration of each trial stayed the same, and the same temporal isolation interval occurred an equal number of times for items at the serial positions 2, 3, and 4. In this way, temporal isolation could be varied independently of serial position, and we controlled the pre-item and post-item intervals separately. The temporal isolation of the first and the last item (list positions 1 and 5) could not be controlled in the same way, and these items were therefore excluded from the analysis of temporal-isolation effects.

**Figure 1.** Experimental procedure



*Figure 1. Example trial.* After a fixation cross displayed for 100 ms on the screen, five colors are presented consecutively, each for 300 ms, separated by inter-item-intervals lasting either 150 ms, 50 ms, 450 ms, or 1350 ms. After presenting the colors, a second fixation cross was presented for 100 ms until a color wheel was displayed on the screen for participants to select the colors they recalled in their order of presentation.

After displaying the last color, a blank screen was shown for 100 ms, followed by a color wheel. The color wheel was rotated randomly in each trial to avoid confounds between colors and their spatial locations in the color wheel. Participants were instructed to reproduce the five colors in the order they had been presented by clicking on the five colors in the color wheel. A new trial began immediately after the last mouse click with a new fixation cross.

### Method: Models

We simulated the experiment with the SOB-CS and the SIMPLE model to obtain predictions. Both models were programmed with MATLAB R2014b. For programming the SOB-CS model, the code of Oberauer, et al. (2012) was used and adjusted to the requirements of this study<sup>1</sup>. The SIMPLE model was programmed based on the equations of Brown, et al. (2007) and extended to the requirements of this study. We next explain how we adjusted and extended each model to apply it to the present experiment. All model features not mentioned here were taken without modification from the published models. For both models we simulated 200 subjects performing the experiment with the same number of trials as the actual participants.

#### SOB-CS

In the current experiment, colors are used as stimuli. The 360 colors participants could select in the color wheel are approximately equidistant values in a circular similarity space. We constructed color representations as vectors of 360 units. Each unit can be thought of as a color-coding neuron with one of the 360 degrees of the color wheel as its preferred color. The tuning function of unit  $i$ , with its preferred color at color-wheel angle  $\theta_i$ , is given by a von Mises distribution with mean  $\mu = \theta_i$  and precision  $\kappa = 4$ . The von-Mises distribution is an approximately normal distribution on the circle. The representation of each color  $i$  is a population code in which each unit is activated according to the angular distance between its preferred color and color  $i$  in the color circle. The activation of each unit in the representation of color  $i$  is given by the von Mises distribution with  $\mu = \theta_i$  and  $\kappa = 4$ . This representational scheme ensures that the similarity between any two colors, measured as the cosine between their two population codes, corresponds to the distance of the two colors in the color wheel.

The value of  $\kappa$  was chosen to generate an overall performance level approximately in accordance with the overall performance level in the experimental data; all other parameter values ( $S_p$ ,  $R$ ,  $r$ ,  $N_o$ ,  $c$ ,  $e$ , and

---

<sup>1</sup> For downloading the MATLAB code of Oberauer, Lewandowsky, Farrell, Jarrold, and Greaves (2012), see [http://www.psychologie.uzh.ch/fachrichtungen/allgpsy/Software\\_en.html](http://www.psychologie.uzh.ch/fachrichtungen/allgpsy/Software_en.html) (22.06.2015)

g) were set to their standard values as in Oberauer et al. (2012). For details on the model and its parameters we refer readers to that publication.

In the experiment each item was presented for 300 ms, followed by a variable inter-item interval. In the model we used the sum of the presentation duration and the following inter-item interval as encoding time. This allows the model to predict a beneficial effect of longer post-item intervals, although that benefit levels off at an encoding time of about 500 ms (Oberauer et al., 2012).

Articulatory suppression is modelled in SOB-CS by encoding the articulated verbal material into working memory. Specifically, each utterance of "bi" results in forming an association between the representation of "bi" and the position of the immediately preceding memory item (Lewandowsky & Oberauer, 2015). Representations of visual stimuli (such as colors) and of syllables (such as "bi") are modelled as only partially overlapping distributed representations (Oberauer et al., 2012). Here, we modelled the representation of each color as a pattern of activation over 360 color-coding units, as described above. The representation of the syllable "bi" was a random vector of +1 and -1 values over 360 units. The 360 units involved in representing colors and the 360 units involved in representing syllables such as "bi" overlap partially; their degree of overlap is controlled by parameter  $p$ , which reflects the proportion of units shared between colors and syllables. Whereas in the representations used in SOB-CS so far, individual units have no specific role and therefore were interchangeable, the representational scheme used for colors assigns each unit to one color (i.e., the preferred color of that unit in the population code). Therefore, we need to carefully select the color-coding units that overlap with the units representing the syllable "bi". To spread the effect of articulatory suppression evenly across representations of all colors, we draw a random subset of the 360 units on each simulation run to overlap with the representation of "bi". Table 1 summarizes all parameters used in the simulation of SOB-CS. The representational overlap parameter  $p$  was set to 0.5, the value used in Simulation 6 of Oberauer et al. (2012).

**Table 1.** Parameters used in SOB-CS

Parameter	Value	Explanation
$s_p$	0.5	Similarity between neighboring positions
$R$	6	Encoding rate
$r$	1.5	Removal rate (for response suppression)
$N_o$	1.5	SD of noise added by output interference
$c$	1.3	Discriminability parameter for translating distance into similarity
$e$	-1000	Threshold for logistic function translating novelty into encoding strength
$g$	0.0033	Gain for logistic function translating novelty into encoding strength
$\kappa$	4	Precision of color representations
$p$	0.5	Proportional overlap between color representations and "bi"

## SIMPLE

SIMPLE assumes better recall of an item with increasing temporal isolation of that item, and with a shorter retention interval (i.e., the temporal distance between the time of retrieval and the to-be-recalled item).

In SIMPLE, temporal distinctiveness of memory representations depends on the temporal distances of all items in memory from the time of retrieval. Temporal distance is measured on a logarithmically compressed psychological time scale. To determine the temporal distance of memory items in each trial at their time of recall, we obtained the time  $t_i$  between encoding (i.e., stimulus presentation) and retrieval (i.e., mouse click on the color wheel) for items in each serial position 1 to 5. We did not record the recall time of each response, but we recorded the overall duration of recall of each trial. On the assumption that recall of each item took about the same time, the time of recall of item  $i$  could be estimated as

$$tr_i = (i - 1) \frac{MR}{5}, \quad (1)$$

where  $MR$  stands for the mean total recall duration across trials ( $MR = 13.2$  s,  $SD = 9.3$  s).

Next we calculated the temporal distance between any two items,  $d(i, j)$ , as:

$$d(i, j) = \text{abs}(\log(t_i) - \log(t_j)). \quad (2)$$

In SIMPLE, there is no established implementation for the difference between the silent and the AS condition. To the extent that colors are represented verbally, and AS disrupts formation or

maintenance of these verbal codes, the memory representations should be less precise under AS. There is no published account of the effect of AS in SIMPLE, so we make the ad-hoc assumption that AS adds noise to all representations in memory, thereby reducing their precision. This can be implemented in SIMPLE by the scaling parameter  $c$ , which controls the general discriminability between memory representations. Therefore, we set  $c$  to a lower value for the AS than for the Silent condition.

The standard version of SIMPLE predicts the probability with which each list item is recalled when a given list position is tested. For forward serial recall, memory is tested first by position 1, then by position 2, and so on. For each tested position, SIMPLE returns a distribution of probabilities  $p(j)$  of recalling list items  $j$ , with  $j$  ranging from 1 to list length  $n$ . Here we need to add a second step at the retrieval process to translate the predicted recall of a list item  $j$  into the predicted probability distribution of selecting each of the 360 colors in the color wheel. We modelled the probability distribution over color space for each list item  $j$  as a von-Mises distribution with the mean set to the true color of item  $j$ , and precision  $\kappa=4$ . The overall probability of selecting each color  $x$  is the sum of the von-Mises distributions centered on each item's true color  $x_j$ , weighted by the probability of recalling list item  $j$ :

$$P(x) = \sum_{j=1}^n p(j) vMises(x, x_j, \kappa). \quad (3)$$

Table 2 summarizes all parameters used in the simulation of SIMPLE. Parameter values for  $c$  were chosen to approximate the overall performance level in the two conditions of the experiment; color precision  $\kappa$  was set to the same value as for SOB-CS.

**Table 2.** Parameters used in SIMPLE

Parameter	Value	Explanation
$c$	[24,12]	Discriminability parameter for translating distance into similarity
$\kappa$	4	Precision of color memory

## Relating Models to Data

We compare the model predictions to the data in two ways. First, we compare model predictions and data with regard to a summary statistic, mean recall error, by experimental condition. Recall error is the absolute deviation between true color and response in degrees on the color wheel. We calculate mean recall error for each serial position, and separately for each level of temporal isolation, to assess the serial-position curves and the effects of pre-item and post-item intervals, respectively. Summary statistics, however, cannot be used to measure transposition errors from continuous-recall data. Therefore, to test the hypothesis that transposition errors obey the "locality

constraint", as found for verbal serial recall, and as predicted by both models, we needed a second approach for comparing the model predictions to the data. We fit the experimental data and the simulation data with a measurement model – the Interference Measurement Model (Oberauer, Stoneking, Wabersich, & Lin, in press) – that estimates the locality constraints directly through one of its parameters.

The Interference Measurement Model (IMM) incorporates the basic principles of cue-based retrieval that underlie both SOB-CS and SIMPLE: The probability of each color to be selected from the color wheel at retrieval depends on the relative strength with which that color is bound to the context that is used as retrieval cue. Applied to serial recall, the context can be either the ordinal serial position or the temporal position of the to-be-recalled item. The color-context binding is maximal between the true color and its true position, and declines gradually with increasing dissimilarity in color space, and with increasing distance in context space. Therefore, a color bound to position  $n$  is also bound to some extent to position  $n-1$  and to position  $n+1$ , and to a smaller extent to positions  $n-2$  and  $n+2$ , and so on. This generalization of color-context bindings along the dimension of contextual distance naturally generates transpositions that obey the locality constraint: When retrieval is cued with position  $n$ , the colors bound to other positions will also be activated to some extent, with their degree of activation declining over positional distance. The steepness of that decline is captured by the contextual-gradient parameter  $s$ . The remaining free parameters of the IMM are the context-independent activation of all colors presented in the current list,  $a$ ; the level of background noise,  $b$ ; and the precision of memory for each list color,  $\kappa$ . The mathematical details of the IMM are presented in the Appendix. We implemented the IMM in a Bayesian hierarchical framework, which provides posterior probability distributions for the group-level parameters (a.k.a. hyper-parameters), representing the mean parameter value in the group.

We fit two versions of the IMM to the experimental data and the simulated data from both models. One version measured contextual distance in terms of ordinal positional distance, incorporating the assumption of event-based models that ordinal list position is the retrieval context used for serial recall. The other version measured contextual distance in terms of (absolute) temporal distance on a logarithmically compressed time dimension, incorporating the assumption of time-based models such as SIMPLE that retrieval of items is cued by their temporal context. By comparing the fit of the two model versions to the empirical data we can obtain further evidence for adjudicating between event-based and time-based models. By fitting both model versions to the simulated data we can verify that the positional version fits the data simulated with SOB-CS better, whereas the temporal version fits the data simulated with SIMPLE better.

## Results

We first present the simulated predictions of both models for the mean error as a function of serial position and temporal isolation, followed by the corresponding experimental results. After that we report the results from fitting the IMM to the experimental and the simulated data.

### **Predictions from SOB and SIMPLE**

Figure 2 show the simulated predictions of SOB-CS, beginning with the mean error as a function of serial position in panel 1, and temporal isolation in panel 2. Panel 3 and 4 depict the separate analysis of pre-item and post-item intervals. The same sequence of predictions for simulations of SIMPLE are shown in Figure 3.

SOB-CS predicts a large primacy effect and no recency effect. The absence of a recency effect is noteworthy because SOB-CS predicts a small recency effect for serial recall of verbal materials (Oberauer et al., 2012). For the present experiment the model predicts no recency effect because the last item was followed by a shorter blank interval (0.1 s) than most other list items, giving it a relatively short encoding time. Overall performance is better in the silent than the AS condition. There is no temporal isolation effect overall. Longer post-item intervals allow longer encoding and thereby improve performance up to about 0.5 s. In the condition with AS the longer post-item intervals also imply longer concurrent articulation, which entails somewhat more interference – this leads to a slight decrement of performance at the longest post-item interval. Longer pre-intervals, in contrast, are not predicted to have any effect.

### **Figure 2.**

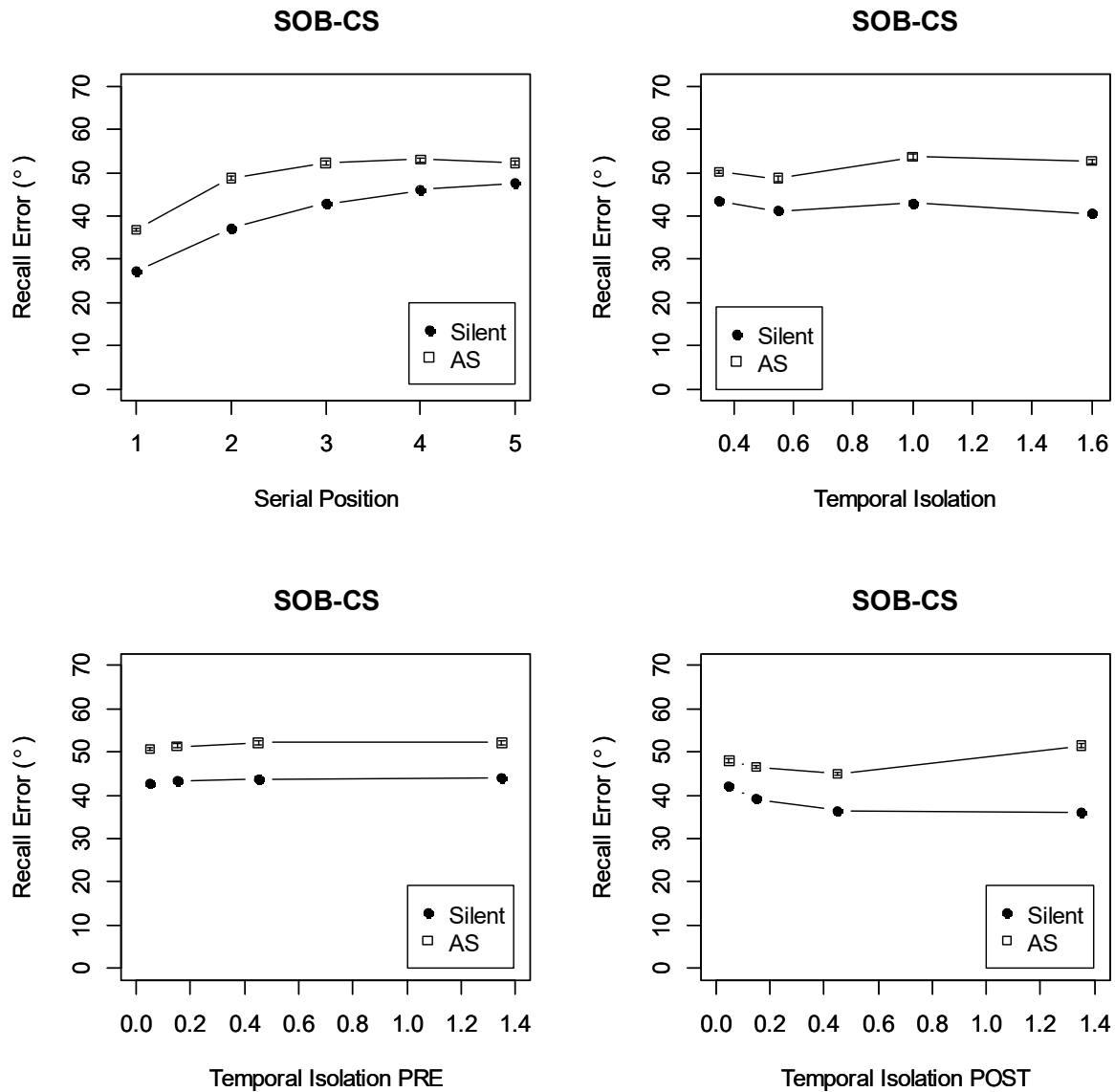


Figure 2. Predictions from simulation with SOB-CS. The first and second panel on the top show the averaged deviation by serial position and by temporal isolation. Panel 3 and 4 show the averaged deviation by pre-item and post-item intervals separately.

SIMPLE predicts a primacy effect and a small recency effect equally for both conditions. It also predicts a strong temporal isolation effect: performance is predicted to improve with increasing temporal isolation intervals. Both longer pre-item and longer post-item intervals are predicted to result in better performance.

Figure 3.



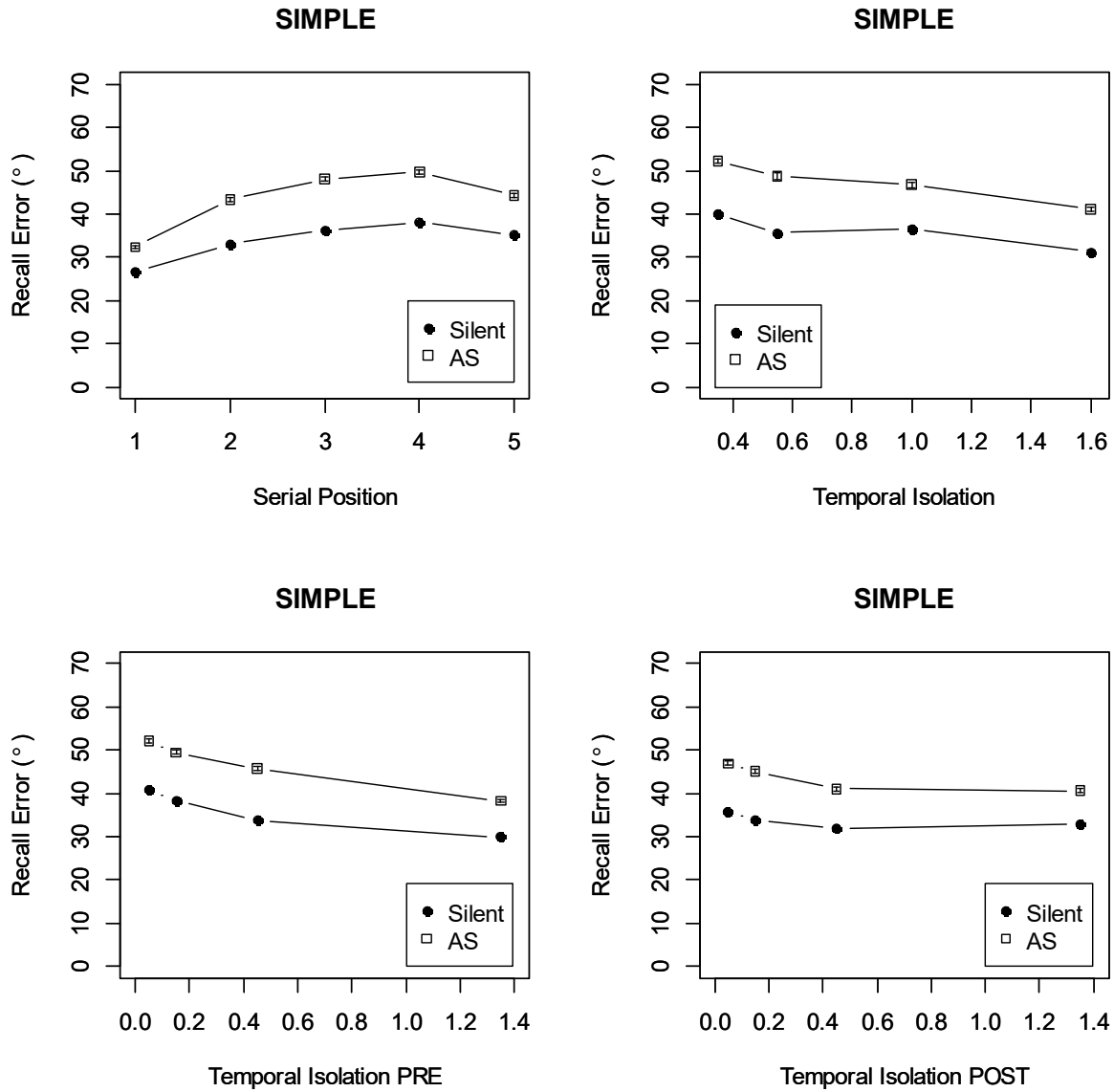


Figure 3. Predictions from simulation with SIMPLE. The first and second panel on the top show the averaged deviation by serial position and by temporal isolation. Panel 3 and 4 show the averaged deviation by pre-item and post-item intervals separately.

## Experimental Results

We analysed the error data with Bayesian hierarchical linear models (Rouder, Morey, Speckman, & Province, 2012) using the *ImBF* function in the BayesFactor package in R (v.0.9.5; Morey & Rouder, 2013). Each analysis started from the full model including the two independent variables (see below) as fixed effects, together with a random main effect of subjects. We assessed the evidence for each main effect and the interaction by comparing the full model to a model with the effect in question removed. The evidence in favor of the effect is expressed as the Bayes factor (BF) for the full model compared to the reduced model. The Bayes factor is the ratio of the marginal likelihoods of the two

models under comparison. The marginal likelihood of a model is the conditional probability of the data, given the model and its parameter values, integrated over all possible parameter values, weighted by the prior probability of these parameter values. We chose the default priors on model parameters (i.e., effect sizes) proposed by Rouder et al. (2012) for all analyses.

For the temporal isolation analysis, temporal isolation (TI) for the serial positions 2, 3, and 4 were computed by adding the pre-item and the post-item interval. The first and the last items were excluded from analysis because these items are separated from other list items by only one temporal interval, and are therefore not comparable to the middle positions. The resulting TI values were combined into four categories to simplify the analysis.

We ran four model comparisons with Bayesian hierarchical linear models. The first set of models included data from all five list positions and tested the effects of serial position (1 to 5) and condition (Silent vs. AS). The second set of models focused on serial positions 2 to 4 and used TI (4 levels) and condition as independent variables. The third set of models tested the effects of pre-item intervals and condition, using data from serial positions 2 to 5. The fourth set of models tested the effects of post-item intervals and condition, using data from serial positions 1 to 4.

As shown in Figure 4, the empirical data show evidence for a serial position effect ( $BF = 10^{32} \times 1.33$ ) in both conditions, silent and articulatory suppression (AS). Earlier stimuli were recalled better than subsequent stimuli, reflecting a strong primacy effect, but there was no recency effect. Performance in the AS condition was worse than performance in the silent condition ( $BF = 10^{23} \times 7.39$ ). There was also evidence for the interaction of these two variables ( $BF = 657$ ).

Table 3 summarizes the results of the model comparisons on the averaged error as a function of temporal isolation and condition (silent vs. AS). The data show a temporal isolation effect in both conditions: Performance improved with larger temporal isolation, with no evidence for an interaction with condition. The final two rows of Table 3 show the model comparisons analysing the effects of pre-item intervals and post-item intervals separately. These analyses provide moderately strong evidence for a beneficial effect of longer post-item intervals for both conditions. The evidence was ambiguous with regard to an effect of pre-item intervals ( $BF$  in favor of the Null hypothesis =  $1/0.8 = 1.25$ ). In both cases there was evidence against the interaction.

Figure 4.

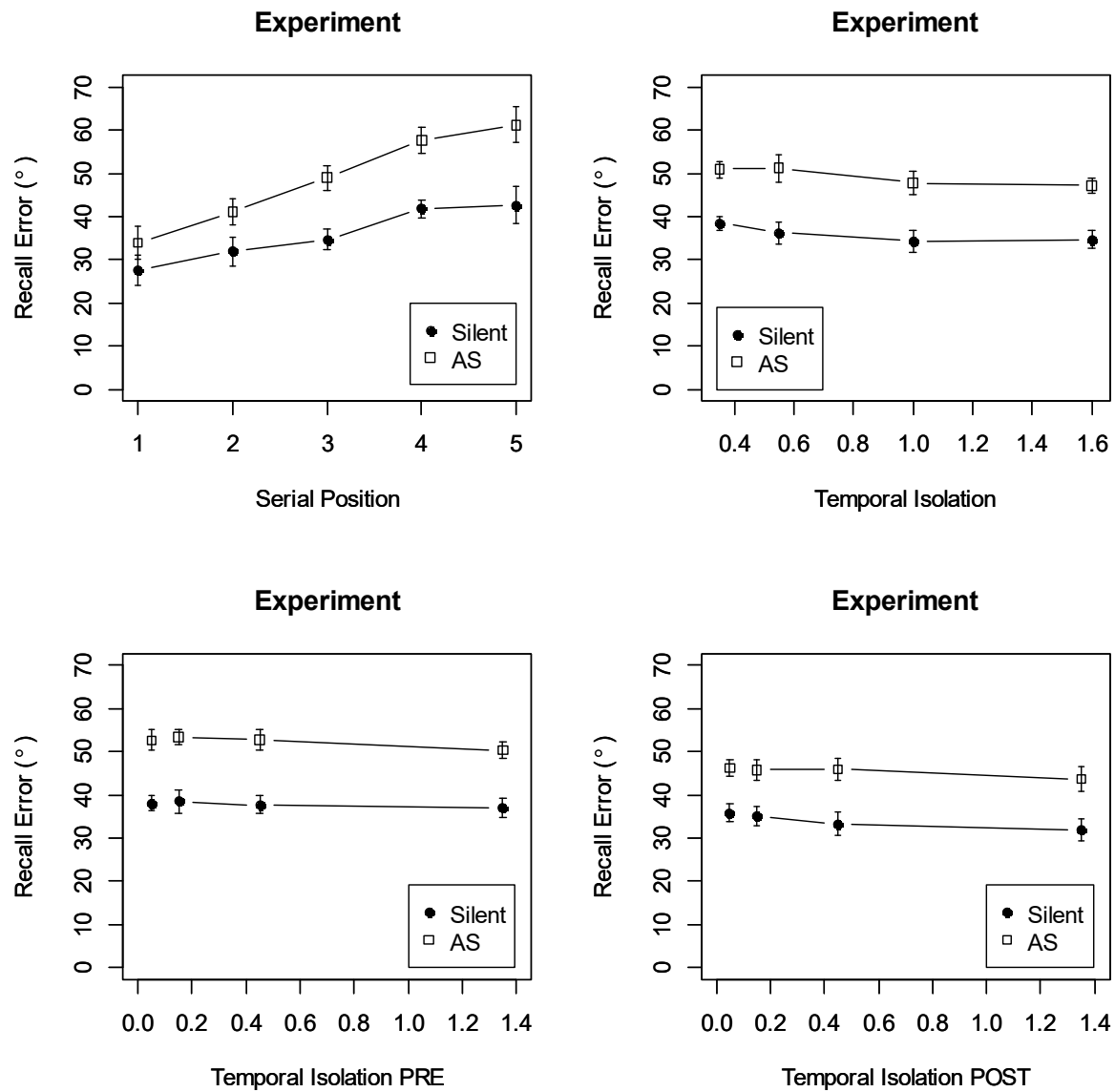


Figure 4. The first and second panel on the top show the averaged deviation by serial position and by temporal isolation. Panel 3 and 4 show the averaged deviation by pre-item and post-item intervals separately.

**Table 3.** Bayes factors of statistical models

Bayes Factors			
IV	Temporal Isolation	Condition (Silent vs. AS)	Interaction
Temporal isolation	1549	$10^{41} \times 2.36$	0.65
Pre-item interval	0.80	$10^{39} \times 1.27$	0.10
Post-item interval	7.16	$10^{25} \times 2.00$	0.12

### Fits with the Interference Measurement Model

We assessed the relative goodness of fit of the two model versions of the IMM – one using ordinal position differences and the other using temporal inter-item intervals as measures of contextual distance – to the data using the Widely Applicable Information Criterion (WAIC; Watanabe, 2010), which is recommended for hierarchical Bayesian models (Gelman, Hwang, & Vetahri, 2014). Lower values reflect a better fit. Table 4 shows the WAIC indices, averaged over serial positions. The experimental data from the Silent condition were fit about equally well by the temporal and the positional model (across the five serial positions, the positional model fit better for the first three positions and the temporal for the last two). The data from the AS condition were fit better by the positional model on average (and for 5 of the 5 serial positions). Although far from conclusive, these results are more in favour of a positional model. As expected, the simulated data from SOB-CS were also fit better by the positional model version, whereas the simulated data from SIMPLE were fit better by the temporal-context version. The model version comparisons for the simulated data confirm that the IMM is able to distinguish between data that were generated from a model using ordinal position representations and a model using temporal context as retrieval cue.

**Table 4.** Model fit statistics (WAIC) for the Interference Measurement Model

Data Source	Condition	Position Version	Temporal Context Version	$\Delta$ WAIC
Experiment	Silent	46332	46334	-2
	AS	57990	58243	- 253
SOB-CS	Silent	110897	112590	-1693
	AS	134413	135548	-1135
SIMPLE	Silent	113498	111810	1687
	AS	117146	116223	1923

Note: Due to memory limitations we computed WAIC for only 50 simulated participants for SOB-CS and SIMPLE.

Figure 5 shows the four free parameters of the IMM (positional-context version) as a function of serial position for the experimental data. Figure 6 presents analogous plots for the simulated data generated by SOB-CS. The parameter estimates from the experimental data differ in two regards from those from the simulated data: First, the effect of serial position was mostly on the precision parameter  $\kappa$  in the data from SOB-CS, whereas it was mostly on the contextual-gradient parameter  $s$  and the background-noise parameter  $b$  in the data. Second, data from SOB-CS resulted in smaller values of the contextual-gradient parameter  $s$ , implying a stronger generalization of positional cueing to neighboring cues. In contrast, the experimental data resulted in larger estimates for parameters  $a$  and  $b$ . Larger values of context-independent activation  $a$  implies more confusions between list colors independent of their positional distance; and larger values of  $b$  imply more responses not reflecting any information about the memory list (i.e., responses driven by random noise or guessing). The difference in the contextual-gradient parameter  $s$  between experimental and simulation data is illustrated in Figures 7 and 8. These figures plot the strength with which each list color was activated at test, given each position as retrieval cue, as predicted by the IMM with parameter values obtained from each data set. The activation of list items in input position  $i$ , given output position  $j$  as retrieval cue, was calculated by the equation:

$$A(i | j) = \hat{a}(i) + \exp[-\hat{s}(j)D(i, j)],$$

with  $\hat{a}$  and  $\hat{s}$  for the IMM parameter estimates (i.e., the means of posteriors of group-level parameters) and  $D(i, j)$  the positional distance between input and output position.

With the parameter values estimated from simulated data, the IMM implies a pronounced positional transposition gradient: Each output position cues not only the item in that position but also the item in neighboring positions, with cueing strength gradually falling off over positional distance. This pattern corresponds to the typical pattern of transposition errors observed in verbal serial recall tasks (e.g., Henson, Norris, Page, & Baddeley, 1996; Lee & Estes, 1977). With the parameter values estimated from the experimental data, each output position cued nearly exclusively the item in that position, with only a hint of a transposition gradient in the later serial positions, somewhat more pronounced with articulatory suppression. Apparently, errors in serial recall of colors with the continuous-reproduction method are much less subject to transpositions as a function of positional distinctiveness than predicted by SOB-CS, and as observed in serial recall of verbal materials.

Figure 5.

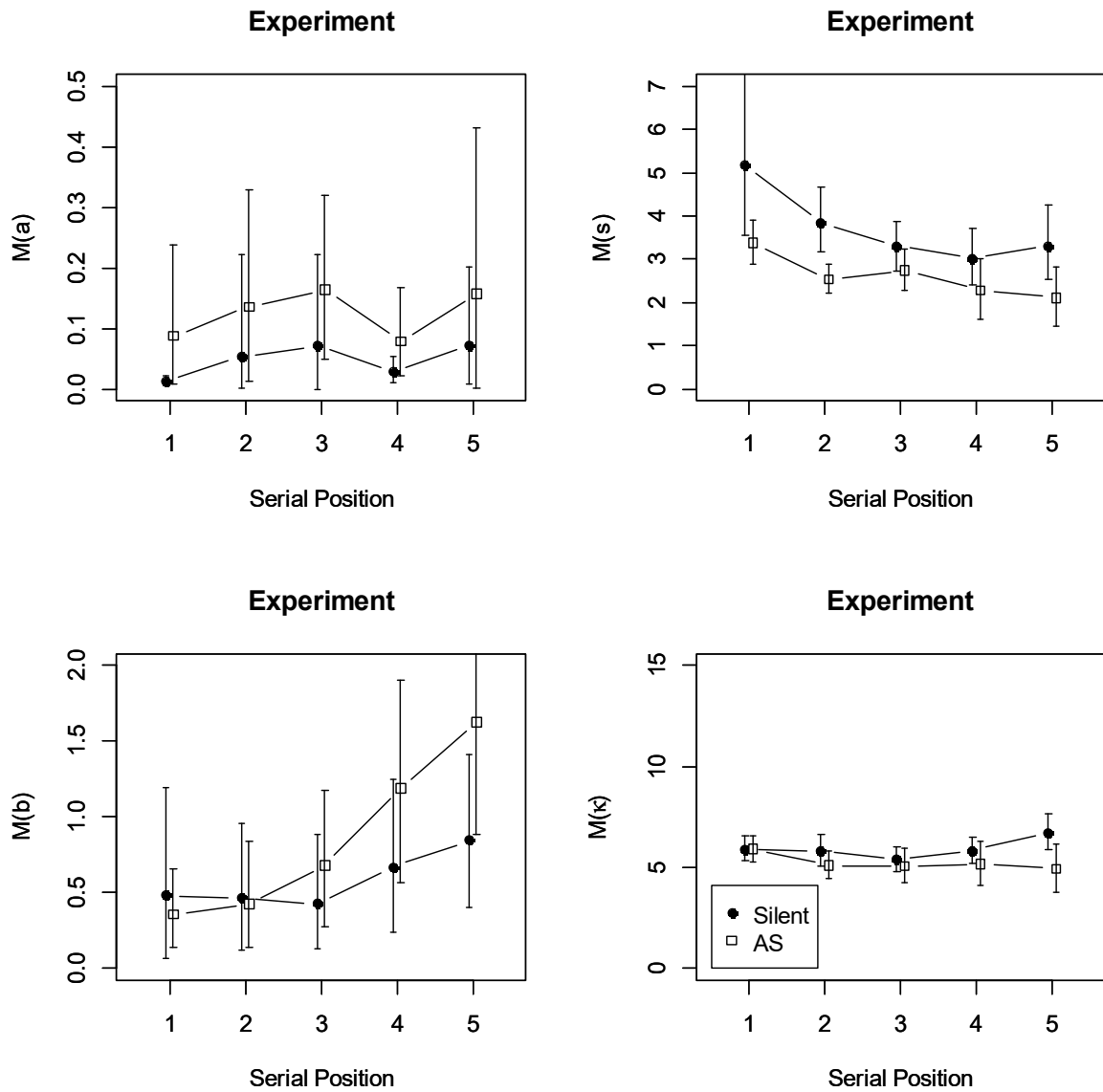


Figure 5. Means and 95% highest-density intervals of the posteriors of the IMM group-level parameters for the experimental data. M(a) is the group mean of context-independent activation of all colors; M(b) the group mean of background noise; M(s) the group mean of the gradient with which cueing strength falls off with positional distance; M(κ) is the group mean of the precision parameter of the von-Mises distribution.

Figure 6.

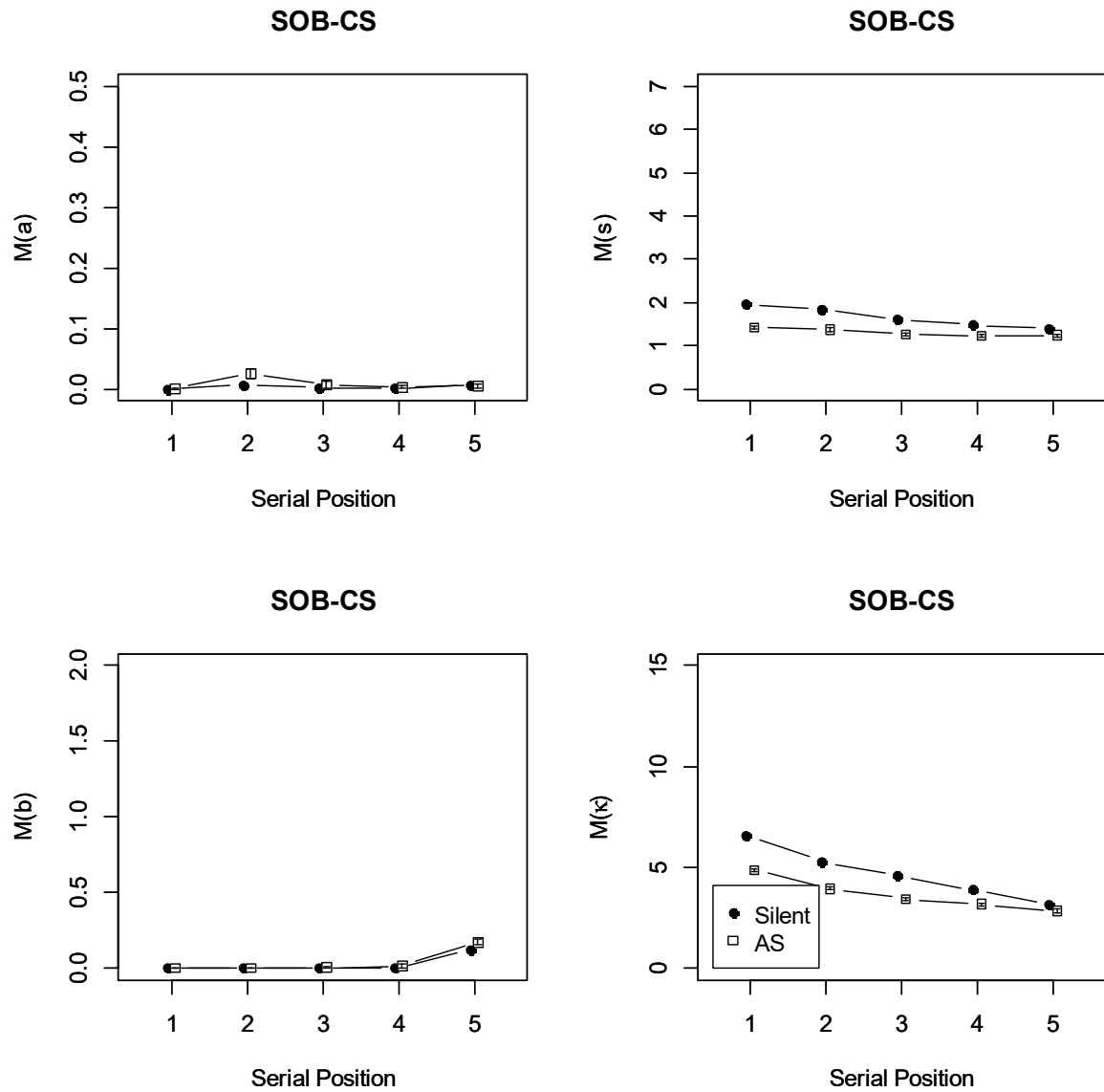


Figure 6. Means and 95% highest-density intervals of the posteriors of the IMM group-level parameters for the data simulated with SOB-CS. M(a) is the group mean of context-independent activation of all colors; M(b) the group mean of background noise; M(s) the group mean of the gradient with which cueing strength falls off with positional distance; M(κ) is the group mean of the precision parameter of the von-Mises distribution.



Figure 7.

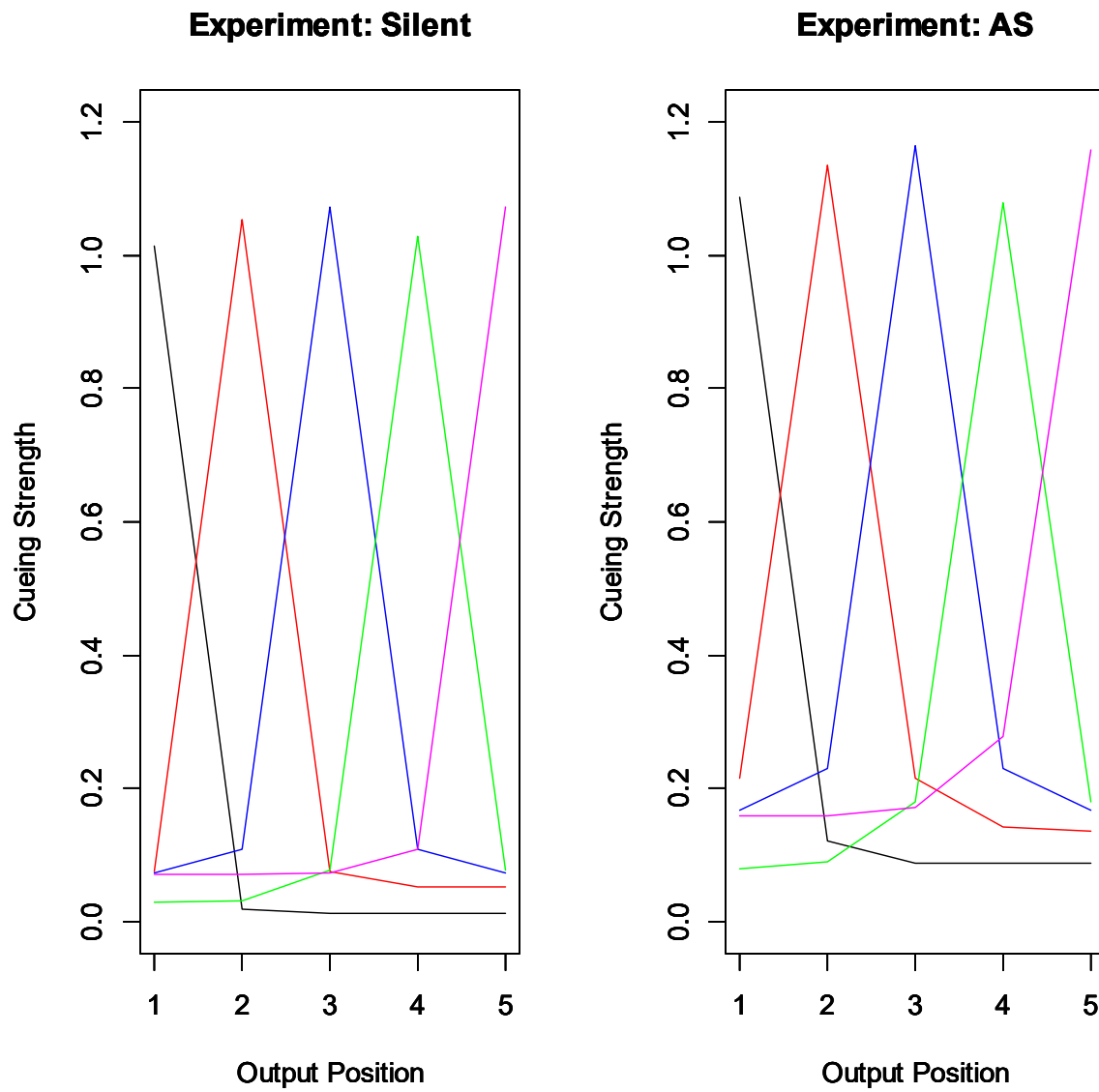


Figure 7: Strength of activation of the color in each list position (line colors, in order black-red-blue-green-magenta) when cued with each output position, as implied by the means of posterior parameter distributions of the IMM: Experimental data.

Figure 8.

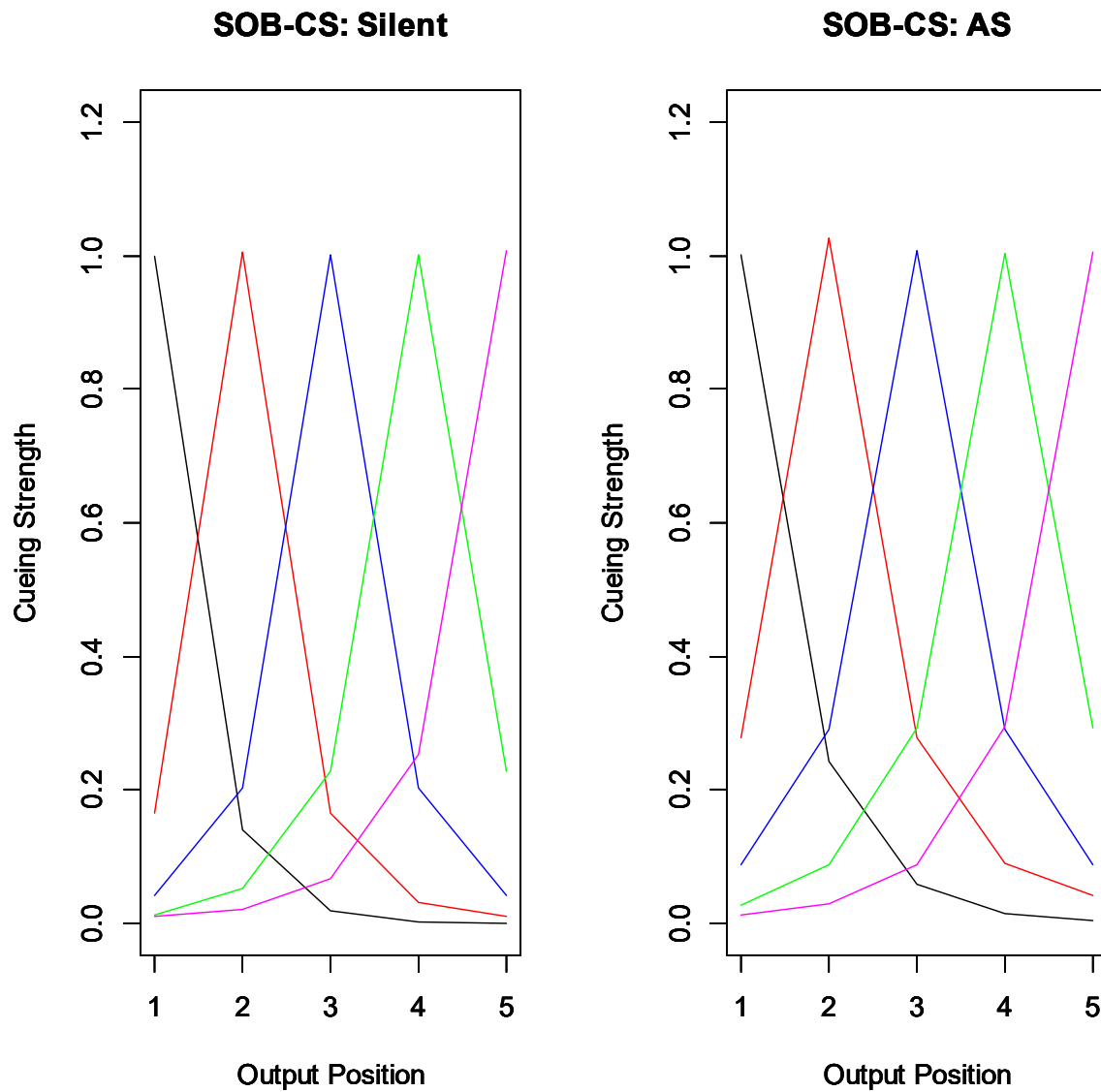


Figure 8: Strength of activation of the color in each list position (line colors, in order black-red-blue-green-magenta) when cued with each output position, as implied by the means of posterior parameter distributions of the IMM: Data simulated from SOB-CS.

## Discussion

We investigated forward serial recall of simple visual stimuli – colors – with the continuous-reproduction paradigm. The empirical data show four findings: first, a strong primacy, but no recency effect. Second, articulatory suppression substantially impaired serial recall of colors. Third, temporal isolation was not universally beneficial – only longer post-item intervals, but not longer pre-item intervals, improved performance. Fourth, serial recall of colors is subject to transpositions that obey the

locality constraint, but the transposition gradient over positional distance appears to be steeper than usually observed with verbal materials.

The finding of strong primacy and no recency effect stands in stark contrast with previous studies of WM for simple visual features using sequential presentation. These studies showed strong recency effects, and at best a small primacy effect (Gorgoraptis et al., 2011; Kool et al., 2014). Our serial-position curves are in better agreement with previous studies using more complex visual stimuli, which observed both primacy and recency effects (Avons, 1998; Avons & Mason, 1999; Smyth et al., 2005; Ward et al., 2005). The key difference between studies that found substantial primacy and those that found predominantly or exclusively a recency effect is that the latter used probed recall, randomly selecting one item to be tested (Gorgoraptis et al., 2011; Kool et al., 2014). In contrast, those studies observing a substantial primacy effect (including ours) tested memory for all items in forward serial order. Comparisons of forward serial recall and probed recall in the verbal domain have shown that the former results in more primacy, and the latter in more recency (Oberauer, 2003). This difference is likely due to two factors. First, recall generates output interference, so that when a list is recalled in forward order, later list items suffer more output interference than earlier items. The SOB-CS model includes output interference as the accumulation of random noise generated by overt recall. In SIMPLE, output interference is a side effect of a loss of temporal distinctiveness over the output sequence when recall takes longer than list presentation, which was the case in our experiment: As recall progresses, the list recedes more and more into the past, reducing the temporal distinctiveness of all list items. Second, when participants know that they will have to recall the list in forward order, they can facilitate that recall order by encoding successive list items with decreasing strength, thereby setting up a primacy gradient of memory strength. In SOB-CS this primacy gradient is a consequence of novelty-gated encoding. In SOB-CS, the primacy gradient is not assumed to be under the person's control. There is evidence, however, that the strength of the primacy effect in list memory is influenced by the person's expectation about the forthcoming test, suggesting that one cause of the primacy effect is a primacy gradient on strength that is at least in part controlled by the person (Duncan & Murdock, 2000; Glanzer & Dolinsky, 1965). A direct comparison of different recall orders with visual materials could determine more precisely the conditions under which the serial-position curve is characterized more by primacy or more by recency.

The pronounced effect of AS converges with Avons (1998), who found a detrimental effect on performance using a serial report task with visual patterns. Additionally, Smyth and colleagues (2005), using a serial reconstruction task with unfamiliar faces as visual stimuli, found a clear decrease of performance through articulatory suppression. Our finding of a pronounced detrimental effect of AS stands in contrast with the well-established absence of any such effect in the change-detection

paradigm for testing visual WM (Morey & Cowan, 2004; Sense, Morey, Prince, Heathcote, & Morey, 2016) ). The change-detection paradigm differs from the present experiment in three important ways: It involves a different kind of test – comparison of a probe to a memory array, rather than reproduction of a memory item – it does not demand memory for serial order, and it involves only a single test. We are not aware of any study comparing silent and articulatory-suppression conditions with the continuous-reproduction task using a simultaneous array with a single test, or with a change-detection task testing memory for serial order, so it is presently unclear whether the form of test (change detection vs. continuous reproduction), the requirement to remember serial order, or the sequential testing of multiple items is responsible for the effect of AS in our experiment. Because of that ambiguity, at present we cannot adjudicate between two explanations for the effect of AS we observed. One explanation is that the representations of colors maintained in WM in our task are not purely visual. It is possible that with sequential presentation and forward serial recall, memory for simple visual features such as colors benefits from an additional verbal representation of the color categories. An alternative explanation, suggested by Smyth et al. (2005), is that AS interferes not with the contents of memory but with the temporal (or positional) context to which they need to be bound to maintain their serial order. If that explanation is correct, the demand on serial-order memory should be the variable that determines whether or not WM for visual information is impaired by AS.

Although we obtained some evidence for a beneficial effect of temporal isolation, this effect was due only to the advantageous effect of longer post-item intervals. Improved memory for items followed by a longer post-item interval is compatible with both time-based and event-based models of serial-order memory. According to the SIMPLE model, both pre-item and post-item intervals are relevant for affecting recall performance, hence a temporal-isolation effect is predicted in both cases. According to the SOB-CS model, longer pre-item intervals should not have a beneficial effect on recall performance – if anything, they should impair memory. Longer post-item intervals could affect recall performance because they provide the opportunity for extended encoding. We used articulatory suppression to control for one such encoding process, the generation of a verbal label. If the beneficial effect of longer post-item intervals was due to the opportunity for creating a verbal representation for the preceding item, then this effect should interact with condition (silent vs. AS). The evidence was unambiguously against such an interaction, ruling out this particular explanation for the post-item interval effect. The beneficial effect of post-item intervals could instead be due to longer time for encoding or consolidation (Ricker & Cowan, 2014) of the visual memory trace.

The beneficial effect of post-item but not pre-item intervals is more compatible with SOB-CS than with SIMPLE. That said, the BF against an effect of pre-item interval was weak, so we cannot rule it out with any confidence. Additional evidence more in favor of SOB-CS than of SIMPLE comes from the

comparison of the two IMM versions: The IMM with positional context – corresponding to event-based models such as SOB-CS – fit the experimental data better than the IMM with temporal context. Taking the evidence from the two analyses together, we tentatively conclude that serial recall of colors is not influenced by temporal distinctiveness. This conclusion is in agreement with that from many studies investigating temporal-distinctiveness effects in serial recall of verbal and spatial materials (Farrell et al., 2011; Lewandowsky et al., 2006, 2008; Nimmo & Lewandowsky, 2005, 2006; Parmentier et al., 2006; Saito, 2000).

Transposition errors are a major source of errors in serial recall of verbal material. These errors are known to obey the locality constraint: List items tend to be confused with other items close by in the list. With the continuous-reproduction paradigm we cannot directly measure transpositions because we cannot unambiguously map responses to list items. Therefore, we tested the hypothesis of a locality constraint on transpositions through a measurement model, the IMM, which estimates the tendency of confusing list items with their positional neighbours through the contextual-gradient parameter. We obtained evidence that transpositions in serial recall of colors obey the locality constraint. At the same time, the transposition gradients implied by the model appear much steeper than those usually found with verbal materials. This observation could help explain the absence of a recency effect in the present data: The small recency effect commonly observed in serial recall is in part due to an "edge effect": The last list item (as well as the first) has neighbors only on one side, reducing its chance of suffering from a transposition error. If transpositions with neighboring list positions are less prevalent in serial recall of colors than of verbal materials, this contribution to the recency effect would be diminished. We note, however, that the comparison of our findings with those from verbal serial recall must be taken with caution because of the many differences between our experiment and our method of measuring the transposition gradient, compared to measures of the transposition gradient in verbal serial recall. A more direct comparison could be made by testing serial recall of visual and verbal materials with a similar method – selection of items from a large set of candidates – and estimating the contextual-gradient parameter of the IMM for both kinds of materials.

To conclude, our investigation of forward serial recall of lists of simple visual features revealed many commonalities with well-established findings from serial recall of discrete stimuli such as colors and spatial locations: The serial-position curve is characterized by pronounced primacy, temporal isolation does not affect performance in the way predicted by temporal-distinctiveness theories, and transposition errors obey the locality constraint. These commonalities suggest that serial recall for visual stimuli obeys largely the same principles as serial recall of other materials (Hurlstone et al., 2014).

## Appendix: The Interference Measurement Model

In the IMM, each retrieval candidate – in the present context, each color on the color wheel – receives a degree of activation at test. The probability of choosing each candidate  $\hat{x}$  out of  $N$  candidates is obtained by Luce's choice rule:

$$P(\hat{x} | P_\theta) = \frac{A(x | P_\theta)}{\sum_{j=1}^N A(j | P_\theta)}$$

Activation arises from three sources: Cue-based retrieval, persistent activation of all colors presented in a list, and background noise.

The activation component from cue-based retrieval,  $A_c$ , reflects the strength with which each color is bound to the currently used retrieval cue  $P_\theta$ . Depending on the model version,  $P_\theta$  indicates the ordinal position of the item to be recalled, or the temporal context of that item. Each color  $i$  of the current list is bound to its position  $P_i$ , and binding strength spreads to neighboring colors on the color wheel according to a von-Mises distribution with mean  $x_i$  (i.e., the presented color) and precision  $\kappa$ . The strength of binding spreads to neighboring positions according to an exponential generalization gradient. Taken together, the activation of each color  $x$  generated through cue-based retrieval with retrieval cue  $P_\theta$  is:

$$A_c(x | P_\theta) = \sum_{i=1}^n \exp[-s \cdot D(P_i, P_\theta)] \cdot vMises(x, x_i, \kappa),$$

The second component,  $A_a$ , is persistent activation of all colors independent of their position, which also spreads to neighboring colors according to a von-Mises distribution. The third component,  $A_b$ , is background noise, which is modeled as a uniform distribution of activation on the circle.

$$A_a(x) = \sum_{i=1}^n vMises(x, x_i, \kappa)$$

$$A_b(x) = \frac{1}{2\pi}$$

The total activation is simply the sum of the three activation components, weighted by parameters  $a$ ,  $b$ , and  $c$ :

$$A(x | P_\theta) = bA_b(x) + aA_a(x) + cA_c(x | P_\theta)$$

Because activation  $A(x|P_\theta)$  is normalized through Luce's choice rule, one of the weighting parameters must be fixed as a scaling parameter for the other two. We fixed  $c = 1$ , so that  $a$  can be interpreted as the strength of context-independent activation of presented items relative to the strength of cue-based retrieval, and  $b$  can be interpreted as the strength of background noise relative to the strength of cue-based retrieval.

We fit the model separately to each serial position in each condition (Silent vs. Articulatory Suppression). We fit the model to data from individual items in each trial, using the hierarchical Bayesian framework developed by Oberauer, Stoneking, Wabersich, and Lin (in press). For reproduction of each item we determined the distance between its position and all other item's positions in the list,  $D(P_\theta, P_i)$ , in terms of ordinal positions (1 to 4) for the position version, and in terms of the temporal-distance calculated according to SIMPLE (our Equation 2) for the temporal-context version. The models were fit to data using JAGS (Plummer, 2016a) and the rjags package for R (Plummer, 2016b).

## References

- Avons, S. E. & Mason, A. (1999). Effects of visual similarity on serial report and item recognition. *The Quarterly Journal of Experimental Psychology: Section A*, 52(1), 217-240.
- Avons, S. E. (1998). Serial report and item recognition of novel visual patterns. *British Journal of Psychology*, 89(2), 285-308.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision*, 10, 433–436.
- Brown, G. D., & Chater, N. (2001). The chronological organization of memory: Common psychological foundations for remembering and timing. *Time and memory: Issues in philosophy and psychology*, 77–110.
- Brown, G. D., Morin, C., & Lewandowsky, S. (2006). Evidence for time-based models of free recall. *Psychonomic Bulletin & Review*, 13(4), 717–723.
- Brown, G. D., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological review*, 114(3), 539.
- Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, 106(3), 551.
- Duncan, M., & Murdock, B. (2000). Recognition and recall with precuing and postcuing. *Journal of Memory and Language*, 42(3), 301-313.
- Farrell, S. (2006). Mixed-list phonological similarity effects in delayed serial recall. *Journal of Memory and Language*, 55(4), 587-600.
- Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic bulletin & review*, 9(1), 59–79.
- Farrell, S., Wise, V., & Lelièvre, A. (2011). Relations between timing, position, and grouping in short-term memory. *Memory & Cognition*, 39, 573-587.
- Geiger, S. M., & Lewandowsky, S. (2008). Temporal isolation does not facilitate forward serial recall—or does it? *Memory & Cognition*, 36(5), 957–967.
- Gelman, A., Hwang, J., & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, 24, 997-1016. doi:10.1007/s11222-013-9416-2
- Glanzer, M., & Dolinsky, R. (1965). The anchor for the serial position curve. *Journal of Verbal Learning and Verbal Behavior*, 4(4), 267-273.
- Glenberg, A. M., & Swanson, N. G. (1986). A temporal distinctiveness theory of recency and modality effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12(1), 3.
- Gorgoraptis, N., Catalao, R. F. G., Bays, P. M., & Husain, M. (2011). Dynamic updating of working memory resources for visual objects. *Journal of Neuroscience*, 31, 8502-8511.



- Henson, R. N. A., Norris, D. G., Page, M. P. A., & Baddeley, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *Quarterly Journal of Experimental Psychology*, 49A, 80-115.
- Hurlstone, M. J., Hitch, G. J., & Baddeley, A. D. (2014). Memory for serial order across domains: An overview of the literature and directions for future research. *Psychological Bulletin*, Advance online publication. doi: 10.1037/a0034221
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, 36(14), 1.
- Kool, W., Conway, A. R. A., & Turk-Browne, N. B. (2014). Sequential dynamics in visual short-term memory. *Attention, Perception & Psychophysics*, 76, 1885-1901.
- Lee, C. L., & Estes, W. K. (1977). Order and position in primary memory for letter strings. *Journal of Verbal Learning & Verbal Behavior*, 16, 395-418.
- Lewandowsky, S., Brown, G. D., Wright, T., & Nimmo, L. M. (2006). Timeless memory: Evidence against temporal distinctiveness models of short-term memory for serial order. *Journal of Memory and Language*, 54(1), 20-38.
- Lewandowsky, S., Duncan, M., & Brown, G. D. (2004). Time does not cause forgetting in short-term serial recall. *Psychonomic Bulletin & Review*, 11(5), 771-790.
- Lewandowsky, S., & Farrell, S. (2008). Short-term memory: New data and a model. *Psychology of Learning and Motivation*, 49, 1-48.
- Farrell, S., & Lewandowsky, S. (2008b). Empirical and theoretical limits on lag recency in free recall. *Psychonomic Bulletin & Review*, 15(6), 1236-1250.
- Lewandowsky, S., & Farrell, S. (2010). *Computational modeling in cognition: Principles and practice*. Sage.
- Lewandowsky, S., Nimmo, L. M., & Brown, G. D. (2008). When temporal isolation benefits memory for serial order. *Journal of Memory and Language*, 58(2), 415-428.
- Lewandowsky, S., & Oberauer, K. (2015). Rehearsal in serial recall: An unworkable solution to the nonexistent problem of decay. *Psychological review*, 122(4), 674.
- Lewandowsky, S., Wright, T., & Brown, G. D. (2007). The interpretation of temporal isolation effects. *The cognitive neuroscience of working memory*, 137-152.
- Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*. doi: 10.1016/j.tics.2013.06.006
- Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience Reviews*. doi: 10.1038/nn.3655
- Morey, C. C., & Cowan, N. (2004). When visual and verbal memories compete: Evidence of cross-domain limits in working memory. *Psychonomic Bulletin & Review*, 11(2), 296-301.

- Morey, R. D., & Rouder, J. N. (2013). BayesFactor (Version 0.9.5). Retrieved from <http://bayesfactorpcl.r-forge.r-project.org/>
- Morin, C., Brown, G. D., & Lewandowsky, S. (2010). Temporal isolation effects in recognition and serial recall. *Memory & cognition*, 38(7), 849–859.
- Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, 18(3), 251–269.
- Nimmo, L. M., & Lewandowsky, S. (2005). From brief gaps to very long pauses: Temporal isolation does not benefit serial recall. *Psychonomic Bulletin & Review*.
- Nimmo, L. M., & Lewandowsky, S. (2006). Distinctiveness revisited: Unpredictable temporal isolation does not benefit short-term serial recall of heard or seen events. *Memory & Cognition*, 34(6), 1368–1375.
- Oberauer, K. (2003). Access to information in working memory: exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3), 411.
- Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: an interference model of complex span. *Psychonomic bulletin & review*, 19(5), 779–819.
- Oberauer, K., Stoneking, C., Wabersich, D., & Lin, H.-Y. (in press). Hierarchical Bayesian measurement models for continuous reproduction of visual features from working memory. *Journal of Vision*.
- Parmentier, F. B., King, S., & Dennis, I. (2006). Local temporal distinctiveness does not benefit auditory verbal and spatial serial recall. *Psychonomic bulletin & review*, 13(3), 458–465.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision*, 10(4), 437–442.
- Plummer, M. (2016a). JAGS 4.2.0. Retrieved from <http://mcmc-jags.sourceforge.net/>
- Plummer, M. (2016b). Package 'rjags', Version 4-6. Retrieved from <http://mcmc-jags.sourceforge.net>
- Ricker, T. J., & Cowan, N. (2014). Differences between presentation methods in working memory procedures: a matter of working memory consolidation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40, 417–428. doi:10.1037/a0034301
- Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, 56(5), 356–374.
- Saito, S. (2000). Effects of articulatory suppression on immediate serial recall of temporally grouped and intonated lists. *Psychologia A*, 22, 142–147.
- Sense, F., Morey, C. C., Prince, M., Heathcote, A., & Morey, R. D. (2016). Opportunity for verbalization does not improve visual change detection performance: a state-trace analysis. *Behavioral Research Methods*. doi:10.3758/s13428-016-0741-1
- Smyth, M. M., Hay, D. C., Hitch, G. J., & Horton, N. J. (2005). Serial position memory in the visual—spatial domain: Reconstructing sequences of unfamiliar faces. *The Quarterly Journal of Experimental Psychology Section A*, 58(5), 909–930.

- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, B*, 64, 583-639.
- Ward, G., Avons, S. E., & Melling, L. (2005). Serial position curves in short-term memory: Functional equivalence across modalities. *Memory*, 13(3-4), 308-317.
- Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and Widely Applicable Information Criterion in singular learning theory. *Journal of Machine Learning Research*, 11, 3571-3594.
- Wills, A. J., & Pothos, E. M. (2012). On the adequacy of current empirical evaluations of formal models of categorization. *Psychological Bulletin*, 138, 102-125. doi:10.1037/a0025715.
- Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of vision*, 4(12), 11-11.